基于门控循环图卷积网络的交通流预测 *

汪 鸣[1]，彭 舰[1†]，黄飞虎[2,1]

(1. 四川大学 计算机学院，成都 610065; 2. 四川中电启明星信息技术有限公司，成都 610041)

摘 要：交通流预测在智能交通系统的建设中起着关键性的作用，然而现有的预测方法无法准确的挖掘其潜在的时空相关性，而且大都采用全连接网络进行单步预测。为了进一步挖掘数据的时空特性以及提升长短期预测的精度，提出了一种门控循环图卷积网络(gated recurrent graph convolutional network，GR-GCN)模型。首先，利用频域上的图卷积结合门控循环单元(GRU)构建一个时空组件(spatial-temporal component，STC)以同时捕获节点的时空相关性，充分的提取数据的时空特征；然后，利用该时空组件构成编码器单元，并将时间序列数据和路网结构数据输入其中；最后，使用门控循环单元作为解码器单元，并按照时间顺序将两者组成一个编码器-解码器(Encoder-Decoder)结构，依次解码出每个时刻的预测结果。在加利福尼亚交通局(Caltrans)性能评估系统中高速公路数据集 PeMSD4 和 PeMSD8 进行了实验。结果表明，所提模型 GR-GCN 在预测未来 15 分钟、30 分钟、45 分钟和 60 分钟的交通流量方面优于大多数现有基准模型，尤其是在长期预测方面。

关键词：交通流量预测；图卷积网络；门控循环单元；编码解码结构

中图分类号：TP183 doi: 10.19734/j.issn.1001-3695.2022.01.0026

Traffic flow prediction based on gated recurrent graph convolutional network

Wang Ming[1], Peng Jian[1†], Huang Feihu[2, 1]

(1. *Sichuan University*, *College of Computer Science*, *Chengdu* 610065, *China*; 2. *Aostar Information Technologies Co*, *Ltd*, *Chengdu* 610041, *China*)

Abstract: Traffic flow prediction plays a key role in the construction of intelligent transportation systems. However, existing prediction methods cannot mine the potential spatiotemporal correlation in the data accurately, and most of them use fully connected networks for single-step prediction. In order to further mine the spatial-temporal features of the data and improve the accuracy of long-term and short-term prediction tasks, this paper proposed a Gated Recurrent Graph Convolutional Network (GR-GCN) . Firstly, by using spectral graph convolution combined with gated recurrent unit (GRU) to construct spatial-temporal components (STC) to capture the spatial-temporal correlation of nodes in the network, fully extract the spatial-temporal features of the data. Secondly, use STC as an encoder unit and enter the time sequence data together with road network data into it. Finally, using gated recurrent unit as a decoder unit, combine the two into an encoder-decoder network structure (Encoder-Decoder) in chronological order, and decode the prediction results at each moment in turn. Experiments were carried out on the highway datasets PeMSD4 and PeMSD8 in the California Department of Transportation (Caltrans) performance evaluation system. The results show that model GR-GCN is better than most existing benchmark models in predicting the traffic flow in the future 15 minutes, 30 minutes, 45 minutes and 60 minutes, especially in long-term prediction.

Key words: traffic flow prediction; graph convolutional network; gated recurrent unit; encoder-decoder architecture

0 引言

随着中国城市化进程的加快，人们对出行的需求日益增大，导致交通拥堵问题日益严重，制约着中国的城市发展。目前亟需建立一套智能交通系统[1]来帮助人们合理规划交通路线以及缓解交通拥堵。而智能交通系统的核心在于准确的预测未来的交通流状况从而辅助系统中的任务。

目前，交通流预测模型主要可以分为三类，第一类为基于统计理论的预测方法，主要包含历史平均分析预测法、自回归差分滑动平均法[2](ARIMA)和卡尔曼滤波分析法[3]。第二类是基于机器学习的方法，此类方法能够建模交通流的非线性状态，且可以通过自适应学习不断调整参数至最优，以此获得更加精确的计算结果，因此逐渐取代统计理论方法。目前用于预测方面的主流机器学习算法包含 KNN[4](K-Nearest Neighbor)、支持向量机[5,6]等。第三类是基于深度学习的方法，也是目前研究的主要方向，深度学习不仅可以学习到数据的特征和相关性，还能在不进行人工特征工程的情况下，从简单特征中挖掘到更为复杂的特征。通过叠加神经网络结构，可以从浅层信息中提取特征，对解决现实生活中的复杂问题有着很大的优势。目前，许多研究者们已经将深度学习运用于交通流量预测领域，并有了显著的成效。代表性的模型有 Liu 等人[7]提出了的 ConvLSTM 模型，该模型结合了 CNN 与 LSTM 来捕获数据的时空相关性。Li 等人[8]提出的 DCRNN 模型利用扩散图卷积对空间特征进行建模，并以基于 GRU(Gated Recurrent Unit)的编码解码结构对时间维度进行建模。Zhao 等人[9]提出了的 T-GCN 模型以及 Yu 等人[10]提出了 STGCN，均用谱图卷积来捕获节点间的空间依赖关系，并利用一维扩张卷积对时间特征进行建模。这些模型通过建模

收稿日期：2022-01-23；**修回日期：**2022-03-07 **基金项目：**四川省重点研发计划(2020YFG0089，22ZDYF3599，2020YFG0304，2020YFG0308)

作者简介：汪鸣(1997-)，男，湖南长沙人，硕士研究生，主要研究方向为图神经网络、时间序列预测；彭舰(1970-)，男(通信作者)，四川成都人，教授，博导，博士，主要研究方向为深度学习、分布式计算(jianpeng@scu.edu.cn)；黄飞虎(1990-)，男，四川遂宁人，讲师，博士，主要研究方向为时空大数据、深度学习.

数据时空相关性来对交通流数据进行预测。但是，以上方法要么没有有效建模数据的时空相关性，无法准确的捕获数据的时空特征，要么就是使用单步预测去预测未来几个时间步的数据，无法准确体现未来交通流数据的时序关系[11]。

交通流量作为一种典型的时空数据，在预测时既需要考虑道路网络的空间依赖性，也需要考虑历史数据以及未来数据的时间依赖性。单独考虑时间或空间依赖，以及笼统的建模未来预测数据之间的关系会丢失对预测结果起作用的重要信息，为解决这一问题，本文提出了门控循环图卷积网络模型(gated recurrent graph convolutional network，GR-GCN)，该模型利用频谱域上的图卷积捕获节点间的空间依赖，并将带有空间关系的节点输入到门控循环单元中同时提取交通流数据的时空特征，然后通过编码解码结构对数据进行多步预测，准确的建模了历史数据间以及未来预测数据间的依赖关系。在真实交通流量预测数据集的实验结果表明本模型效果优于基准模型。

1　门控循环图卷积网络

1.1　交通流预测任务

交通流预测任务是一个典型的时间序列预测任务，本文的研究只涉及交通流数据，因此只用保留交通状态数据中的一个属性，即 $\boldsymbol{X}\in\boldsymbol{R}^{N\times T\times 1}$ 。交通流预测问题是根据给定历史观测数据 $\boldsymbol{X}=\{\boldsymbol{x}_{t-h+1},\boldsymbol{x}_{t-h+2},\ldots,\boldsymbol{x}_t\}\in\boldsymbol{R}^{N\times H\times 1}$ 和对应时刻的路网结构 $\boldsymbol{G}$ ，其中 H 为历史序列输入的窗口大小，预测未来 T_p 时段所有节点的交通流数据 $\boldsymbol{Y}=\{\boldsymbol{y}_{t+1},\boldsymbol{y}_{t+2},\ldots\boldsymbol{y}_{t+T_p}\}\in\boldsymbol{R}^{N\times T_p}$ ，其中 $\boldsymbol{y}_{t+T_i}=\{y^1_{t+T_i},y^2_{t+T_i},\ldots y^N_{t+T_i}\}\in\boldsymbol{R}^N$ 表示在未来 T_i 时间段中整个交通网络的流量状态。简而言之，交通预测问题的本质是学习一个函数 F ，根据历史观测数据和对应的路网结构 $\boldsymbol{G}$ ，预测未来时段 T_p 内所有节点的交通流数据 $\boldsymbol{Y}$ ，具体的公式如下：

$$(\boldsymbol{y}_{t+1},\boldsymbol{y}_{t+2},\ldots\boldsymbol{y}_{t+T_p})=F[(\boldsymbol{x}_{t-h+1},\boldsymbol{x}_{t-h+2},\ldots\boldsymbol{x}_t),\boldsymbol{G}] \tag{1}$$

1.2　图卷积网络

卷积神经网络是以卷积操作为基础的前馈神经网络，可以高效的处理图像、语音这类欧式数据。但是实际生活中有很多数据都是非欧式数据，这类数据并不能直接使用卷积网络直接进行特征提取，因此提出了图卷积网络[12,13]，其是一种将卷积操作扩展到图结构的神经网络，核心思想是中心节点对其邻域节点进行信息聚合。表 1 展示了进行图卷积操作涉及的一些符号定义。

表 1　图相关定义

Tab. 1　The definition of graph

符号	含义
G	图
E	图中边集合
V	图中节点的集合
A	图的节点邻接矩阵
D	图中节点的度矩阵
L	图的拉普拉斯矩阵
I	单位矩阵
$H^{(l)}$	网络中第 l 层的特征

其中，邻接矩阵反映着图中节点之间相邻关系，度矩阵反映了与每个节点相连的边数量。最能反映图结构性质的是图的拉普拉斯矩阵，该矩阵是图研究领域中的核心研究对象，描述的是中心节点与邻居节点信号的差异，可以反映节点邻域的平滑性，具体定义为

$$\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{A} \tag{2}$$

通常，正则化的拉普拉斯矩阵更为常用，定义为

$$\boldsymbol{L}^{sym}=\boldsymbol{D}^{-1/2}\boldsymbol{L}\boldsymbol{D}^{-1/2}=\boldsymbol{I}-\boldsymbol{D}^{-1/2}\boldsymbol{A}\boldsymbol{D}^{-1/2} \tag{3}$$

GCN 模型的原理是利用卷积本质上是频域上的滤波这一特性，在频域上进行操作[14]。将节点特征看成信号，通过傅里叶变换到频域空间，最后通过傅里叶逆变换得到图域卷积。每一层卷积仅处理一阶邻域信息，通过叠加若干卷积层可以实现多阶邻域的信息传递[15]。每一个卷积层的传播规则如下：

$$\boldsymbol{H}^{(l+1)}=\sigma(\boldsymbol{L}^{sym}\boldsymbol{H}^{(l)}\boldsymbol{W}^{(l)}) \tag{4}$$

其中：$\sigma(.)$ 表示非线性激活函数。简而言之，GCN 的每一层通过邻接矩阵和特征矩阵相乘得到每个顶点邻居特征的汇总，然后再乘上一个参数矩阵，最后用激活函数得到聚合邻居节点特征的矩阵。一个两层的图卷积网络模型可以表示为如下：

$$f(\boldsymbol{X},\boldsymbol{A})=\sigma(\hat{\boldsymbol{A}}ReLU(\hat{\boldsymbol{A}}\boldsymbol{X}\boldsymbol{W}_0)\boldsymbol{W}_1) \tag{5}$$

其中：$\boldsymbol{X}$ 表示特征矩阵，$\boldsymbol{A}$ 表示拓扑网络的邻接矩阵，$\hat{\boldsymbol{A}}$ 表示使用拉普拉斯正则化后的邻接矩阵，$\boldsymbol{W}_0$ 和 $\boldsymbol{W}_1$ 分别为第一层与第二层的权重参数，$\sigma(.)$ 和 $ReLU(.)$ 表示网络所用的激活函数。

1.3　门控循环单元

门控循环单元[16]网络(Gated Recurrent Unit，GRU)是基于循环神经网络(Recurrent Neural Networks，RNN)的改进模型。相比于 RNN，它能够解决长期记忆和反向传播中的梯度消失等问题。相比于长短期记忆网络(Long-Short Term Memory，LSTM)，使用 GRU 能够与其达到相当的效果，并且相比之下更容易进行训练，能够很大程度上提高训练效率。

GRU 中的隐藏单元是一个特殊的细胞结构而不是传统神经网络的节点。在其内部有两个门控，即一个重置门(reset gate)和一个更新门(update gate)。直观上说，重置门决定了如何将新的输入信息与前面的记忆相结合，更新门决定了前面记忆保存到当前时间步的量。GRU 的细胞结构如图 1 所示。

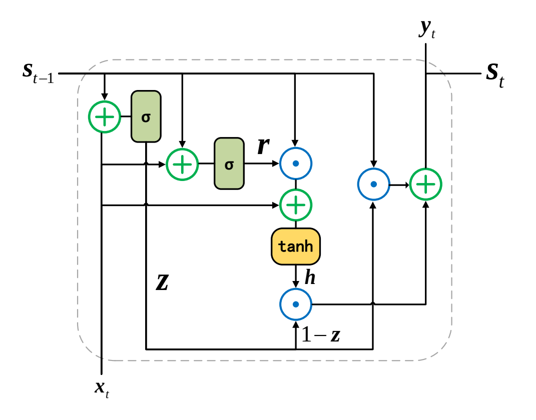


图 1　GRU 单元结构

Fig. 1　The inner structure of GRU

各个门控的计算公式如下：

$$\begin{aligned}\boldsymbol{z}&=\sigma(\boldsymbol{x}_t\boldsymbol{U}^z+\boldsymbol{s}_{t-1}\boldsymbol{W}^z)\\ \boldsymbol{r}&=\sigma(\boldsymbol{x}_t\boldsymbol{U}^r+\boldsymbol{s}_{t-1}\boldsymbol{W}^r)\\ \boldsymbol{h}&=tanh(\boldsymbol{x}_t\boldsymbol{U}^h+(\boldsymbol{s}_{t-1}\odot\boldsymbol{r})\boldsymbol{W}^h)\\ \boldsymbol{s}_t&=(1-\boldsymbol{z})\odot\boldsymbol{h}+\boldsymbol{z}\odot\boldsymbol{s}_{t-1}\end{aligned} \tag{6}$$

其中：符号 "⊙" 表示两个向量对应位置上的元素相乘；σ 表示 sigmoid 激活函数；$\boldsymbol{z}$ 代表更新门；$\boldsymbol{r}$ 代表重置门；$\boldsymbol{h}$ 代表当前记忆内容；$\boldsymbol{s}_t$ 代表当前时间步的最终记忆。由于其具有循环结构，上一个细胞的最终记忆和当前时间步的输入一起作为当前细胞的输入。通过串联这些单元结构，并将数据按照时间顺序输入到这些单元中，经过这些单元的运作和学习，即能捕捉到数据中潜在的时间相关性。

1.4　编码解码结构

编码器-解码器模型(Seq2Seq)最初是用于机器翻译领域的一种很成功的结构[17]。这种模型接受一个序列作为输入，并将序列中的信息编码为中间表示，最后使用解码器解码中

间表示依次解码为目标序列，由于使用解码器来预测部分，预测结果一般不可知，往往需要将前一时刻的预测结果和编码的隐藏表示一起作为下一时刻的输入。如图 2 所示。

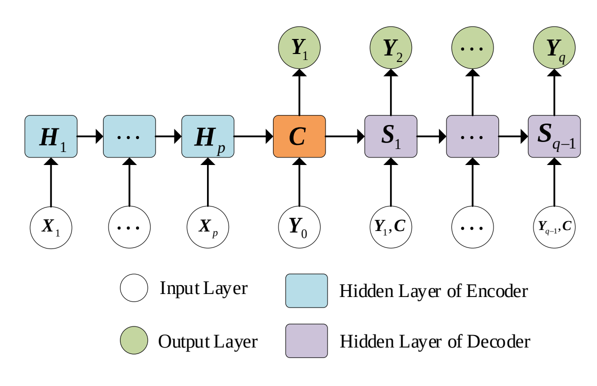


图 2 编码器解码器结构

Fig. 2 The structure of Encoder-Decoder

这类结构具有很强的变通性，其编码器和解码器可以是不同类型的结构，在现有的研究中，编码器可以为 RNN(recurrent neural network)、GRU(gated recurrent units)、LSTM(long short term memory)、GCN(graph convolutional network)或者时空注意块(spatial-temporal attention block)的任意组合,同时解码器可以为 RNN、GRU、LSTM 或者多层感知机的任意组合[18,19]。只用满足如下映射关系：

$$X=[X_1,\cdots,X_p]\xrightarrow{seq2seq}Y=[Y_1,\cdots,Y_q] \tag{7}$$

其中：p 与 q 分别为编码器和解码器的长度。这两者可以不相等，因此可以将其应用于交通预测任务中，对数据进行多步预测。具体来说，编码器将历史序列编码为一个隐藏中间表示 $\boldsymbol{C}$，然后将中间隐藏表示 $\boldsymbol{C}$ 输入到解码层使其解码成为未来的交通状态。

1.5 门控循环图卷积网络模型

为了同时捕获交通流数据的空间和时间的相关性，本文基于图卷积网络、门控融合单元以及编码器解码器提出了一个门控循环图神经网络模型，如图 3 所示。

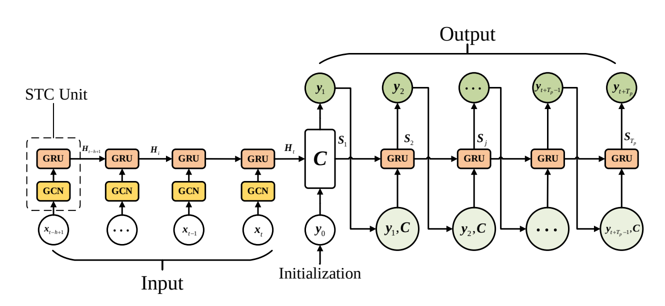


图 3 GR-GCN 整体模型结构图

Fig. 3 The overall architecture of GR-GCN

图 3 的左半部分属于编码器，将历史序列数据按照时间步依次输入到编码器的各个单元中，本文将此类单元称之为时间图卷积单元。单元运作过程如下，首先，单元中的图卷积网络被用来捕获交通拓扑路网的结构并得到其空间特征。然后，将带有空间信息的时间序列数据输入到门控循环单元，信息在这些单元中动态变化来得到序列数据的时间特征。当整个历史序列输入完并处理完成后，最后一个单元的输出即为中间表示 $\boldsymbol{C}$。编码器的过程如下：

$$\begin{gathered}\boldsymbol{H}_i=GRU[GC(\boldsymbol{x}_i),\boldsymbol{H}_{i-1}]\\ \boldsymbol{C}=\boldsymbol{H}_t,\boldsymbol{S}_0=\boldsymbol{H}_t\end{gathered} \tag{8}$$

其中：$\boldsymbol{H}_i$ 是与输入 $\boldsymbol{x}_i$ 相关联的隐藏表示，$\boldsymbol{H}_0$ 是随机初始化的非零矩阵。

图 3 的右半部分属于解码器，解码器则使用 GRU 单元对中间表示 $\boldsymbol{C}$ 进行解码，解码后得到隐藏表示。通过一个前馈神经网络对未来某个时间步的数据进行预测。各个时间步的隐藏状态及预测值 $\hat{y}_j$ 的具体计算方式如下：

$$\begin{gathered}\boldsymbol{S}_j=GRU(\boldsymbol{C},\hat{\boldsymbol{y}}_{j-1},\boldsymbol{S}_{j-1})\\ \hat{\boldsymbol{y}}_j=\boldsymbol{S}_j\boldsymbol{W}\end{gathered} \tag{9}$$

其中：$\boldsymbol{S}_j$ 是与输出 $\hat{y}_j$ 相关联的隐藏表示，并且 $\hat{y}_0$ 是输出序列的开始，属于一个序列的初始信号，需要进行初始化，在模型中使用上一个时间步的流量特征对其进行初始化，便于解码器的迭代。

对于时空建模单元(TGCN)，其内部结构如图 4 所示。

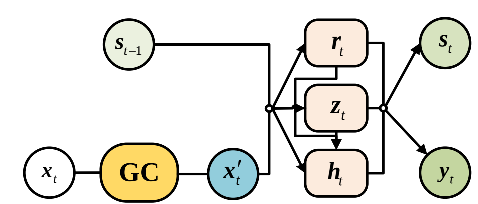


图 4 时间图卷积单元处理过程

Fig. 4 The process of temporal graph convolution unit

图中，$\boldsymbol{s}_{t-1}$ 表示在 $t-1$ 时刻的输出，$\boldsymbol{GC}$ 代表图卷积处理过程，$\boldsymbol{r}_t$ 与 $\boldsymbol{z}_t$ 分别代表在 t 时刻的重置门与更新门，$\boldsymbol{s}_t$ 表示在 t 时刻的输出，$\boldsymbol{y}_t$ 表示 $\boldsymbol{s}_t$ 经过全连接层变换的输出。具体的计算过程如下：

$$\begin{gathered}\boldsymbol{z}_t=\sigma(\boldsymbol{W}^z[f(\boldsymbol{A},\boldsymbol{x}_t),\boldsymbol{s}_{t-1}]+\boldsymbol{b}_z)\\ \boldsymbol{r}_t=\sigma(\boldsymbol{W}^r[f(\boldsymbol{A},\boldsymbol{x}_t),\boldsymbol{s}_{t-1}]+\boldsymbol{b}_r)\\ \boldsymbol{h}_t=tanh(\boldsymbol{W}^h[f(\boldsymbol{A},\boldsymbol{x}_t),(\boldsymbol{s}_{t-1}\odot\boldsymbol{r}_t)]+\boldsymbol{b}_h)\\ \boldsymbol{s}_t=(1-\boldsymbol{z}_t)\odot\boldsymbol{h}_t+\boldsymbol{z}_t\odot\boldsymbol{s}_{t-1}\end{gathered} \tag{10}$$

其中：$f(\boldsymbol{A},\boldsymbol{x}_t)$ 表示图卷积的处理过程，计算方式如式(5)，$\boldsymbol{W}$ 和 $\boldsymbol{b}$ 表示训练过程中的权重系数与偏置。

解码过程与编码过程类似，由于在解码过程中，输入数据已经经过编码器中的图卷积进行了空间依赖性捕获，则不需要再对其进行图卷积操作了。因此，在解码器部分，本文仅使用 GRU 对中间变量进行解码，解码方式如式(6)。

总的来说，时空图卷积网络模型能够处理交通数据中复杂的空间依赖性和动态的时间依赖性。一方面，图卷积网络被用来捕获交通城市网络拓扑结构中节点之间的空间依赖关系。另一方面，门控循环单元被用来捕获交通道路上信息的动态变化以得到节点自身的时间依赖性。最后使用编码器解码器(Seq2Seq)结构完成对交通流的多步预测。通过以上模块的有效结合，很好的建模了交通流的空间和时间方面的依赖性，并完成预测任务。

2 实验结果与分析

2.1 数据集

本文利用两个真实的高速公路数据集 PeMSD4 和 PeMSD8 进行实验。使用其中的车流量数据作为实验数据集，使用采集频率为 30 秒/次的原始数据汇总成时间间隔为 5 分钟的数据样本，即每个点一个小时包含 12 个交通数据样本。

本文使用过去一个小时的流量数据(12 个监测样本)去预测未来 15 分钟(3 个样本)、30 分钟(6 个样本)、45 分钟(9 个样本)和 60 分钟(12 个样本)的流量数据，并按照 7 : 1 : 2 的比例将数据集划分成为训练集、验证集和测试集，并且利用训练集的均值和方差对划分后的三个数据集进行归一化处理，数据的详细信息可见表 2。

2.2 对比模型和评价指标

实验通过两个回归任务常用的指标来评估模型的性能，分别为平均绝对误差(Mean Absolute Error，MAE)均方根误差(Root Mean Squared Error，RMSE)，具体的计算公式为

$$MAE=\frac{1}{m}\sum_{i=1}^{m}\left|Y_i-\hat{Y}_i\right| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{m}^{1}(Y_i - \hat{Y}_i)^2} \tag{12}$$

本文选择结合了 MSE 和 MAE 的 Huber 作为损失函数，通过调节超参数来确定 MAE 和 MSE 的作用范围，能够使得模型尽快收敛到全局最优解，具体的计算公式为

$$L_\delta(Y,\hat{Y}) = \begin{cases} \frac{1}{2}(Y-\hat{Y})^2 & if\ |Y-\hat{Y}| \le \delta \\ \delta|Y-\hat{Y}| - \frac{1}{2}\delta^2 & others \end{cases} \tag{13}$$

表 2　数据集描述

Tab. 2　The description of data sets

数据集	PeMSD4	PeMSD8
传感器数量	307	170
时间范围	1/1/2018 – 2/28/2018	7/1/2016 – 8/31/2016
样本间隔	5 分钟	
训练样本数	11877	12482
验证样本数	1696	1783
测试样本数	3392	3566

为了验证本文提出模型的有效性，设置了 7 种基准方法进行实验对比。

HA：历史平均模型。HA 模型直接使用平均值作为预测结果，本文使用 12 个历史时间片的平均值来预测未来若干个值。

ARIMA[2]：自回归差分滑动平均模型。ARIMA 模型时应用较为广泛的时间序列分析模型。

GCN[20]：图卷积网络。将时间维度作为图中节点的特征，最后使用全连接进行预测任务。

GRU[16]：门控循环单元网络。GRU 是一种特殊的循环神经网络模型，实验中使用一层 GRU 和一层全连接的模型组合。

T-GCN[21]：时间图卷积网络。T-GCN 将图卷积集成与循环神经网络单元中，并用一层全连接的模型进行预测。

DCRNN[8]：扩散图卷积循环网络。这是一个基于 GNN 和 RNN 的模型，它集成了 GRU 和双向扩散卷积用于预测。

AGCRN[22]：自适应图卷积循环网络。它采用自适应图并将 GRU 与图卷积与节点自适应参数学习相结合，并使用全连接进行预测。

2.3　实验超参设置与分析

本文模型 GR-GCN 主要涉及的超参数包括学习率(learning rate)、训练轮数(epoch)、批大小(batch size)和隐藏层数量。本文设置学习率为 0.001，batch 大小设置为 64、训练轮数设置为 200 轮。

神经网络隐藏单元的数量是本文模型重要的一个参数，不同的隐藏单元的数量会很大程度的影响模型的预测精度。为了能够选择出最适合的值，本文设计了相关实验。

在 PeMSD4 数据集上，隐藏单元数量分别设置为 8、16、32、64 与 128 并且分析了预测误差的变化。如图 5 所示，横坐标表示隐藏单元的数量，纵坐标表示预测误差。

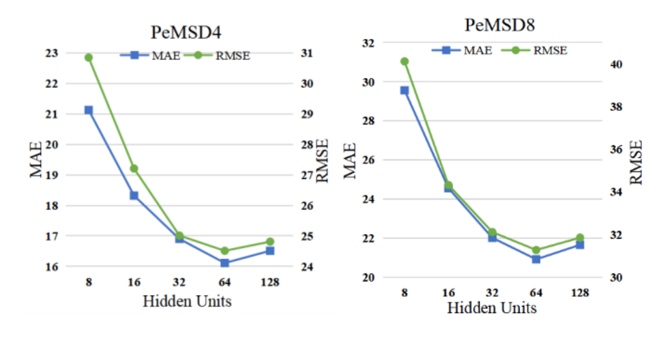


图 5　在不同隐藏单元数量下模型的预测误差

Fig. 5　The prediction error of proposed model under different number of hidden units

图 5 反映了在 PeMSD4 与 PeMSD8 数据集下随着隐藏单元数量的变化，平均绝对误差 MAE 和均方误差 RMSE 的大小。可以观察到，随着隐藏单元数量增大，误差是先增大后减小的。当隐藏单元数量为 64 时，模型的预测误差最小。主要是因为当隐藏单元逐渐增大到一定的值后，模型的复杂度以及计算复杂度会呈指数型增大，因此，本文在整个实验过程中，将隐藏层数均设置为 64。

实验及结果分析

表 3 展示了本文模型和基准模型在 PeMSD4 和 PeMSD8 数据集上预测 15 分钟、30 分钟、45 分钟与 60 分钟时交通流量预测的误差表现。

表 3　各个模型在不同预测范围上的误差表现

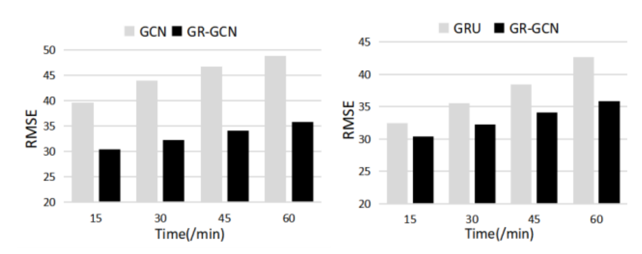
Tab. 3　The performance of proposed model and baselines on PeMSD4 and PeMSD8

Data	Method	15min MAE	15min RMSE	30min MAE	30min RMSE	45min MAE	45min RMSE	60min MAE	60min RMSE
PeMSD4	HA	28.22	41.93	31.59	46.82	34.95	51.78	38.29	56.76
	ARIMA	24.36	43.47	30.59	58.30	37.62	67.86	44.31	74.89
	GCN	24.66	36.16	26.24	39.92	28.91	45.66	31.70	50.12
	GRU	23.36	33.84	24.58	37.81	26.64	40.15	28.55	42.53
	T-GCN	21.01	31.94	23.90	34.69	26.69	37.09	30.29	40.28
	DCRNN	20.45	30.65	22.43	32.67	24.89	34.90	26.54	36.98
	AGCRN	**19.83**	**29.85**	**21.46**	**31.81**	23.74	33.45	25.04	35.16
	GR-GCN	20.81	30.97	21.97	32.11	**23.10**	**33.09**	**24.54**	**34.41**
PeMSD8	HA	23.08	34.26	26.05	38.66	29.02	43.08	31.97	47.49
	ARIMA	18.25	29.58	22.96	39.61	28.16	50.03	33.87	60.15
	GCN	19.87	28.34	21.41	34.77	23.57	39.96	24.87	44.40
	GRU	16.62	24.61	19.32	29.74	21.18	32.11	23.07	35.71
	T-GCN	17.83	25.53	18.09	27.36	20.14	29.96	21.18	31.47
	DCRNN	16.36	24.02	17.25	26.61	18.38	27.35	19.63	28.56
	AGCRN	15.56	**22.86**	16.72	24.42	17.31	25.80	18.22	27.19
	GR-GCN	**15.26**	23.31	**16.29**	**24.01**	**16.68**	**25.12**	**17.82**	**26.34**

从上表可以看出，就两个评估指标而言，本章提出的模型在两个数据集上都接近最佳表现。本节将从三个方面来分析了模型 GR-GCN 的优势。

首先是拥有低误差优势，可以发现，一些基于神经网络，并且充分考虑时间特征建模重要性的模型，包括 GR-GCN、T-GCN 以及 GRU，普遍在预测精度方面会高于其他像 HA、ARIMA 这类基于统计理论的预测模型。以数据集 PeMSD4 上对于提前 15 分钟的预测任务为例，GR-GCN 和 T-GCN 的均方误差比 HA 的均方误差分别低了 27.5%和 14.1%，比 ARIMA 的均方误差分别低了 30.0%和 17.1%。这是因为像 HA、ARIMA 这类模型无法处理时间序列这种复杂、非线性的数据。而 GCN 之所以具有很高的误差是由于 GCN 只考虑了空间特性而忽略了交通数据是一个典型的时间序列数据。除此之外，ARIMA 作为一个比较成熟的时间序列预测模型，预测的准确性比 HA 模型更低的原因在于其难于处理长时以及非平稳的序列数据，并且在计算误差时使用的是总体误差的一个平均值，如果某个节点数据波动较大，则使得最终的预测效果变差。

接着是模型的时空预测能力优势，为了验证模型时空预测方面的优势，本文对比了 GCN 和 GRU 这种神经网络模型。从图 6 可以清晰的看到基于时空特征的模型(GR-GCN)比基于单个特征(GCN、GRU)进行预测效果要更好，以数据集上 PeMSD4 上对于提前 15min 的预测任务为例，GR-GCN 的均方误差相比 GCN 的均方误差低了 32.5%，比 GRU 的均方误差低了 11%。说明 GR-GCN 模型能够很好的捕获数据中的时间和空间相关性



(a) GCN 与 GR-GCN 误差对比　(b) GRU 与 GR-GCN 误差对比

图 6　GCN 和 GRU 与 GR-GCN 误差对比图

Fig. 6　The error comparison among GCN, GRU and GR-GCN

最后就是对长时预测能力优势。为了更直观的说明此方面的优势，绘制了根据预测步长的变化，预测误差变化趋势图，如图 7 与图 8 所示。从图中可以看出，随着预测步数的增大，其他基准模型的均方误差比本文模型 GR-GCN 增长速度更快，例如，在 PeMSD4 上预测单个时间步时，T-GCN 与 GRU 的预测误差跟 GR-GCN 基本持平，但是在预测 12 个时间步时，差距明显增大，GR-GCN 比同样为时空预测模型的 T-GCN 误差小了 11.1%。除此之外，相较于一些在短期预测上比较与优秀的模型，如 DCRNN 和 AGCRN，本文所提模型在短期预测方面(小于 30min)有所不足，均方误差比 AGCRN 大约高了 3%-5%。但是，在长期预测方面，本文模型 GR-GCN 得益于其编码-解码的预测结构，较这两者模型表现略好，GR-GCN 在预测 60min 时比 AGCRN 的均方误差小了 3.1%。这证明了模型的编码器-解码器架构的加入，对数据的长期预测时有一定效果的。

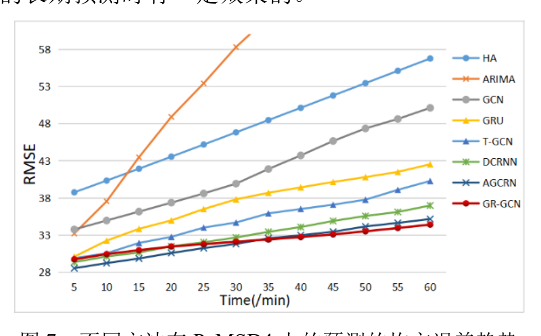


图 7　不同方法在 PeMSD4 上的预测的均方误差趋势

Fig. 7　The RMSE trend on pemsd4 by different methods

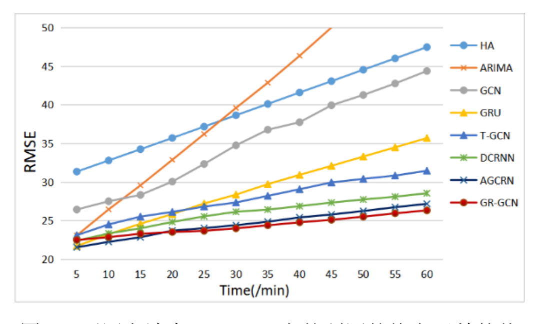


图 8　不同方法在 PeMSD8 上的预测的均方误差趋势

Fig. 8　The RMSE trend on pemsd8 by different methods

为了能更好的理解 GR-GCN 模型的预测效果，在实验结束环节，本文在 PeMSD4 上随机选择了两个不同的路段，对这两个路段提前 15 分钟和提前 60 分钟的交通流量预测情况进行可视化，对比预测值与真实值之间的差距，如图 9 和 10 所示。其中灰色虚线代表真实的数据，绿色曲线代表提前 15 分钟的预测结果，红色曲线代表提前 60 分钟的预测结果。

2.4　模型消融实验

此外，本文对模型 GR-GCN 进行了消融实验(Ablation Experiment)，操作方法类似于控制变量法，目的是为了证明模型当中的各个子模块对于预测能力的提升能力。分别去除图卷积模块(Our-GC)和解码器模块(Our-De)对其进行分析。在 PeMSD4 和 PEMSD8 数据集上观察两种变体模型和模型 GR-GCN 分别在 15 分钟和 60 分钟时平均绝对误差(MAE)的情况，结果如图 11 所示。

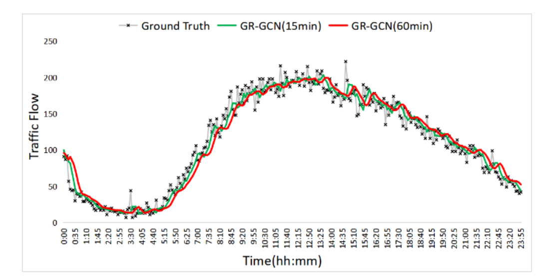


图 9　模型在传感器 1 上预测情况

Fig. 9　The comparison of predictions on sensor 1

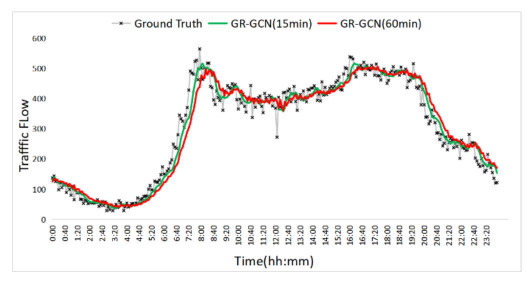
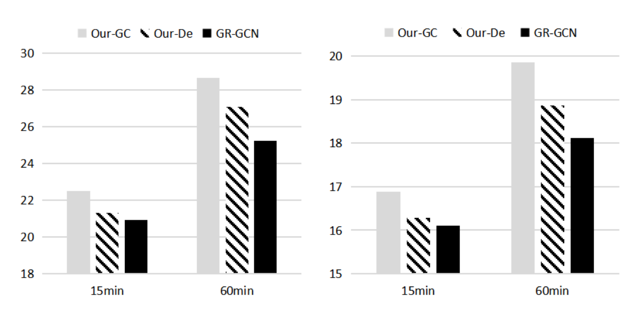


图 10　模型在传感器 2 上预测情况

Fig. 10　The comparison of predictions on sensor 2



(a) PeMSD4 消融实验误差对比　(b) PeMSD8 消融实验误差对比

图 11　消融实验误差对比图

Fig. 11　The MAE of the ablation experiments

从上图可以看出，在模型去除图卷积网络的情况下，无论是短期(15min)预测任务还是长期(60min)预测任务，预测误差都有着明显的增大。这再次说明，同时考虑交通数据的时间和空间的特征比仅仅考虑时间维度的特征会更有利于预测。对于去除解码器模块，改用全连接进行单步预测的情况，在短期预测任务下误差与本文所提模型没有较大差距，但是对于长期预测任务差距明显，这再次说明，使用能够进行多步预测的解码器是有利于长期预测的。

综上所述，本文所提模型的各个模块不仅对短期预测任务有较好的提升效果，还有利于降低在长期预测任务上的预测误差。

3　结束语

本文针对现有的交通流预测方法未能有效的结合时空特征以及大部分使用全连接层进行单步预测的问题。提出了基于编码器解码器结构的神经网络模型。该模型在编码器部分时空 GCN 提取相邻节点交通流量的空间信息，再将带有空间信息的节点通过 GRU 提取节点中的时序信息，并得到一个包含时空特征的隐藏表示。最后通过解码器依次解码出每个时间步的预测信息，得出预测结果。最终在两个真实路网数据集上的实验结果表明，模型在 MAE 和 RMSE 两个指标值下基本上优于其他基线算法，可以有效的对交通流量进行预测。

参考文献：

[1] 徐红兵，邓惠俊. 人工智能技术在城市智能交通系统中的应用 [J]. 计算机产品与流通，2020 (01): 167. (Xu Hongbing, Deng Huijun. Application of atrificial intelligence technology in urban intelligent transportation system [J]. Computer Products and Circulation, 2020 (01): 167)

[2] Williams B M, Hoel L A. MODELING AND FORECASTING VEHICULAR TRAFFIC FLOW AS A SEASONAL STOCHASTIC TIME SERIES PROCESS [J]. Dissertation Abstracts International, Volume: 60-01, Section: B, page: 0292. ; Adviser: Lester A. Hoel, 1999.

[3] Okutani I, Stephanedes Y J. Dynamic prediction of traffic volume through Kalman filtering theory [J]. Transportation Research Part B Methodological, 1984, 18 (1): 1-11.

[4] Lint H V, Hinsbergen C V. Short-Term Traffic and Travel Time Prediction Models [J]. Transportation Research E-Circular, 2012.

[5] Drucker H, Burges C J C, Kaufman L, *et al.* Support Vector Regression Machines [J]. Advances in Neural Information Processing Systems, 1997, 28 (7): 779-784.

[6] 刘钊，杜威，闫冬梅，等. 基于K近邻算法和支持向量回归组合的短时交通流预测 [J]. 公路交通科技，2017, 34 (05): 122-128+158. (Liu Zhao, Du Wei, Yan Dongmei, *et al.* Short-term traffic flow prediction based on combination of K-nearest neighbor algorithm and support vector regression [J]. Journal of Highway and Transportation Research and Development, 2017, 34 (05): 122-128+158.)

[7] Liu Yipeng, Zheng Haifeng, Feng Xinxin, *et al.* Short-term traffic flow prediction with Conv-LSTM [C]. 2017 9th International Conference on Wireless Communications and Signal Processing (WCSP) , 2017: 1-6.

[8] Li Yaguang, Yu R, Shahabi C, *et al.* Diffusion convolutional recurrent neural network: Data-driven traffic forecasting [J]. arXiv preprint arXiv: 1707. 01926, 2017.

[9] Bai Jiandong, Zhu Jiawei, Song Yujiao, *et al.* A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting [J]. International Journal of Geo-Information, 2021, 10 (7): 485.

[10] Yu Bing, Yin Haiteng, Zhu Zhanxing. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting [J]. arXiv preprint arXiv: 1709. 04875, 2017.

[11] Ye Jiexia, Zhao Juanjuan, Ye Kejiang, *et al.* How to Build a Graph-Based Deep Learning Architecture in Traffic Domain: A Survey [J] , 2020.

[12] Niepert M, Ahmed M, Kutzkov K. Learning convolutional neural networks for graphs [C]. International conference on machine learning, 2016: 2014-2023.

[13] Wu Zonghan, Pan Shirui, Chen Fengwen, *et al.* A comprehensive survey on graph neural networks [J]. IEEE Trans on neural networks and learning systems, 2020, 32 (1): 4-24.

[14] Bruna J, Zaremba W, Szlam A, *et al.* Spectral networks and locally connected networks on graphs [J]. arXiv preprint arXiv: 1312. 6203, 2013.

[15] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering [J]. Advances in neural information processing systems, 2016, 29: 3844-3852.

[16] Chung J, Gulcehre C, Cho K H, *et al.* Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling [J]. Eprint Arxiv, 2014.

[17] Luong M T, Le Q V, Sutskever I, *et al.* Multi-task Sequence to Sequence Learning [J]. Computer Science, 2015.

[18] Lu Mingming, Zhang Kunfang, Liu Haiying, *et al.* Graph Hierarchical Convolutional Recurrent Neural Network (GHCRNN) for Vehicle Condition Prediction [J] , 2019.

[19] Chen Weiqi, Chen Ling, Xie Yu, *et al.* Multi-Range Attentive Bicomponent Graph Convolutional Network for Traffic Forecasting [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34 (4): 3529-3536.

[20] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks [J]. arXiv preprint arXiv: 1609. 02907, 2016.

[21] Zhao Ling, Song Yujiao, Zhang Chao, *et al.* T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction [J]. IEEE Trans on Intelligent Transportation Systems, 2019, PP (99): 1-11.

[22] Bai Lei, Yao Lina, Li Can, *et al.* Adaptive Graph Convolutional Recurrent Network for Traffic Forecasting [J] , Advances in Neural Information Processing Systems 2020, 33: 17804-17815.